Домашнее задание к лекции 2019-10-08 по теме "Вероятностные основы линейной регрессии". вес = 0.2, максимальный балл = 100.

На лекции была выведена функция ошибки $MSE(y_i, x_i)$ как следствие заранее предположений:

- L (linearity)
- I (Independence)
- N (normal distribution of error)
- E (equivariance of error)

Было показано, что все эти предпосылки приводят к тому, что наибольшее правдоподобие выборки $p(y_i, x_i|\theta)$ достигается при значении параметров, оптимизированных при решении задачи минимизации среднеквадратического отклонения на обучающей выборке:

$$\hat{\theta} = \underset{\Theta}{\operatorname{argmin}} \left(\sum_{i=1}^{i=N} (y_i - \theta \bar{x_i})^2 \right)$$

Напомним, что такая форма функции потерь продиктована распределением ошибки, которое предполагается нормальным (предположение "N"):

$$y_i = \theta \bar{x_i} + \epsilon,$$

$$\epsilon \sim \mathcal{N}(0, \sigma^2),$$

то есть,

$$p(\epsilon) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\epsilon^2}{2\sigma^2} \right).$$

При этом принцип максимизации правдоподобия - отправная точка во всем выводе.

Однако нужно понимать, что предположение о нормальности ошибки может недостаточно точно описывать имеющиеся данные. Исходя из каких-то сторонних соображений, исследователь может принять решение о том, что ошибка для каждого x_i распределена иным образом, и именно из этого распределения генерируется реализация y_i случайной величины Y в точке x_i - то есть, для объекта x_i, описываемого вектором признаков $\bar{x_i}$

Задание: в предположении, что ошибка распределена по Лапласу, вывести новый вид функции потерь:

$$\epsilon \sim Laplace(0, \alpha)$$

$$p(\epsilon) = \frac{1}{2\alpha} \exp\left(-\frac{|\epsilon|}{\alpha} \right)$$